Amino Terminal Sequence and Location of Phosphate Groups of the Major Human Casein

RAE GREENBERG, MERTON L. GROVES, and ROBERT F. PETERSON
Eastern Regional Research Center[1]
Philadelphia, PA 19118

ABSTRACT

The amino acid sequence of the first 28 residues of the major human casein was determined. This protein in multiphosphorylated forms (0 to 5 phosphorous per molecule) was compared to cow beta-casein which is similar in composition but phosphorylated at a constant level. After sequencing the phosphate-free human casein, phosphorylated seryl and threonyl residues were located in three of the other phosphorylated forms by examining the aqueous layer of the phenylthiohydantoin conversion step during automatic liquid phase sequencing. Phosphate groups on specific seryl/threonyl residues suggest a biosynthetic mechanism involving stepwise phosphorylation or dephosphorylation.

INTRODUCTION

The major component of the casein system in mature human milk is a protein which occurs in multiphosphorylated forms containing zero to five phosphate groups per moelcule (6, 13). It is similar in amino acid composition and electrophoretic mobility to cow (*Bos taurus*) β-casein, a protein of constant phosphate content of four or five per molecule (*M* 24,000) (15), depending upon the genetic type. The polymorph designated β-casein C is the one containing only four phosphate groups (18). Human casein and cow β-casein each contain 210 amino acid residues per molecule and in composition differ by only 35 residues. A comparison of the sequences and phosphorylation patterns of the human and cow proteins is of interest in connection with studies of the evolutionary divergence of mammalian species (8).

This report presents the sequence of the first 28 amino acid residues (13%) of the major human casein and the locations of phosphate groups on three of the forms phosphorylated at different levels.

MATERIALS AND METHODS

Whole casein from individual mature human milks (milk after the 10th day of lactation) was prepared as described previously by Groves and Gordon (6). Before chromatography, casein solutions were heated at 95 C for 5 min to eliminate any possible proteolysis during the fractionation procedure. Separation and purification of the components were carred out on DE32-microgranular cellulose by stepwise column chromatography in phosphate buffer (.005 to .10 M) at pH 8.3 (6). Purity and identity of the fractions were ascertained by disc gel electrophoresis (pH 9.6, 4 M urea). Six pure proteins having identical amino acid compositions but differing in phosphate content from zero to five residues per molecule were isolated and four of these, the zero, bi-, tetra-, and penta-phosphorylated forms, were used for sequence studies.

Amino acid sequence determination was on a Beckman 890C[2] sequencer with a Quadrol double cleavage program. Identification of the phenylthiohydantoin (PTH) amino acids was accomplished by gas and thin-layer chromatography (14, 9) and/or amino acid analysis after back hydrolysis with HI to the parent amino acid after Smithies et al. (16).

During this investigation, it was established that the phosphorus-containing derivative produced from a phosphoseryl or phosphothreonyl residue by an Edman cycle remains in the aqueous layer during the PTH conversion step. This was confirmed by a detailed study of the phosphorylated region of cow β-casein, a phosphoprotein of known sequence (15). Thus,

Received November 10, 1975.

[1] Agricultural Research Service, US Department of Agriculture.

[2] Reference to brand or firm name does not constitute endorsement by the US Department of Agriculture over others of a similar nature not mentioned.

after ascertaining the sequence of the nonphosphorylated human casein, phosphate groups were positioned in the P-containing forms by microphosphorus analysis (12) of the aqueous layer at each serine or threonine site. At times, on duplicate runs, the conversion step was omitted and the sequencer product in regions of interest examined directly for phosphorus. This approach also proved successful. Although the sequencer yield decreases after passing through an area containing phosphoserine, continued identification of sequential residues is still possible. Basing the calculation on yields determined before and after such an area, phosphorus recoveries averaged better than 90% of theoretical.

RESULTS AND DISCUSSION

Figure 1 shows the results of amino terminal sequence analysis on the intact phosphate-free form of the major human casein. Residue 13, which remains unidentified, must be reexamined when fragments of the molecule are sequenced. The biphosphorylated form exhibited the same sequence with phosphorylated seryl residues at positions 9 and 10. Serines 6, 8, 9, and 10 were phosphorylated in the tetra-P form. The threonyl residue at position 3 as well as the aforementioned serines is phosphorylated in the penta-P molecule. The finding of phosphorus only on specific serine and threonine residues in forms phosphorylated at different levels indicates that these forms are indeed homogeneous with respect to phosphoserine sites and not merely charge class mixtures averaging two or four phosphate groups with random phosphorylation.

The first 18 amino acids are similar in composition to a phosphopeptide isolated from a tryptic hydrolysate of whole human casein by Strid (17) long before the casein system had been well characterized. A comparison of the N-terminal human casein and cow β-casein sequences is shown in Fig. 2. Matching up the phosphorylated regions of the two proteins maximizes the homologies. This requires placing the amino terminal residue of the human protein at position 10 of the cow sequence as shown in Fig. 2 or alternatively allowing for a gap of nine residues after matching up positions 1 and 2. Since both proteins contain the same number of residues, one might speculate that a compensating extension or insertion will occur later in human casein. Further comparison must await the completion of the human casein amino acid sequence.

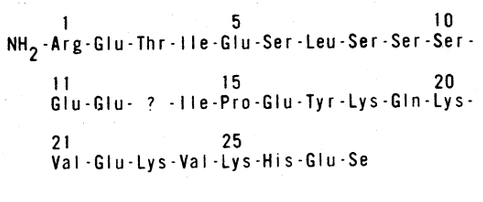
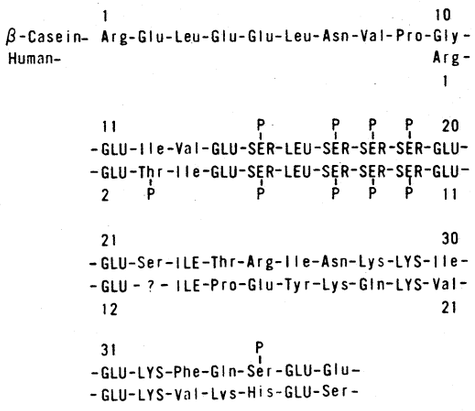


Figure 2 also shows the identity of the phosphorus-containing segments of the two proteins: cf., human casein (residues 5 to 12) and cow β-casein (residues 14 to 21). The same sequence cluster also is found in the phosphorylated region of α_{s1}-casein (residues 63 to 70) (11). Mano and Imahori (10), working with phosvitins, suggested that during phosphorylation there is selective incorporation of phosphate into phosphoserine blocks and that a neighboring phosphate group or acidic amino acid promotes further incorporation of a phosphate group. Mercier et al. (11) have theorized that phosphorylation in the cow caseins occurs when the requirement Thr/Ser-X-Glu/PSer (where X is any amino acid) is met. The



penta-phosphate form of human casein also seems to follow this pattern. The location of phosphorus on seryl residues 9 and 10 in the biphosphorylated form leads to the speculation that position 10 would be phosphorylated first, this being mediated by the highly acidic nature of positions 11 and 12 (-Glu-Glu-). Stepwise phosphorylation back to serine 6 then might occur. The threonyl residue at position 3 (see Fig. 1) is followed by a glutamic acid at position 5 and is phosphorylated as postulated by the aforementioned theory.

Present knowledge of casein biosynthesis in the lactating mammary gland may be summarized as follows: casein is synthesized on the polyribosomes of the endoplasmic reticulum (5); the nascent polypeptide chains are phosphorylated after synthesis from a pool of unphosphorylated casein (19); the subcellular sitc of phosphorylation is probably the Golgi vesicles (2); the casein subunits are packaged into micelles in the Golgi vacuoles (4) and thereby transported into the alveolar lumen (7). Bingham and Farrell (2, 3) have isolated a casein kinase from the Golgi apparatus of rat and cow lactating mammary gland which phosphorylates dephosphorylated caseins. The cow caseins as isolated do not exhibit the multiphosphorylation phenomenon; each one manifests a constant phosphate content. The obvious and still unanswered question concerns the difference in human casein biosynthesis which allows the "leakage" of unphosphorylated or partially phosphorylated forms of the protein into the secreted milk. With present knowledge, enzymatic dephosphorylation cannot be disregarded as an equally plausible explanation. This also would have to be a specific stepwise process and would have to occur after full phosphorylation in the Golgi apparatus. An acid phosphatase active on phosphoprotein substrates has been characterized in cow milk by Andrews and Pallivicini (1).

Studies are underway to elucidate the complete primary structure of the major human casein and to examine its biosynthesis possibly involving stepwise phosphorylation or dephosphorylation.

2 Bingham, E. W., and H. M. Farrell, Jr. 1974. Casein kinase from the Golgi apparatus of lactating mammary gland. J. Biol. Chem. 249:3647.
3 Bingham, E. W., and H. M. Farrell, Jr. 1974. Casein kinase: A Golgi apparatus enzyme from the lactating mamary gland of cow. Fed. Proc. 33:1381.
4 Farrell, H. M., Jr. 1973. Models for casein micelle formation. J. Dairy Sci. 56:1195.
5 Gaye, P., L. Houdebine, and R. Denamur. 1973. Isolation of active messenger RNA for α_s-casein from bound polyribosomes of mammary gland. Biochem. Biophys. Res. Commun. 51:637.
6 Groves, M. L., and W. G. Gordon. 1970. The major component of human casein: A protein phosphorylated at different levels. Arch. Biochem. Biophys. 140:47.
7 Heald, C. W., and R. G. Saacke. 1972. Cytological comparison of milk protein synthesis of rat mammary tissue in vivo and in vitro. J. Dairy Sci. 55:621.
8 Jenness, R. 1974. Page 3 *in* Lactation: A comprehensive treatise. Vol. 3. B. L. Larson and V. R. Smith, ed. Academic Press, New York.
9 Jeppsson, J. O., and J. Sjöquist. 1967. Thin-layer chromatography of PTH amino acids. Anal. Biochem. 18:264.
10 Mano, Y., and K. Imahori. 1970. Structure of phosphoprotein metabolically active in phosphorylation and dephosphorylation reactions. J. Biochem. (Tokyo) 67:767.
11 Mercier, J.-C., F. Grosclaude, and B. Ribadeau-Dumas. 1971. Structure primaire de la caseine α_{s1}-bovine-Sequence complète. Eur. J. Biochem. 23:41.
12 Meun, D. H. C., and K. C. Smith. 1968. A micro phosphate method. Anal. Biochem. 26:364.
13 Nagasawa, T., I. Kiyosawa, and K. Kuwahara. 1970. Human casein. II. Isolation of human β-casein fraction and human β-casein B. J. Dairy Sci. 53:136.
14 Pisano, J. J., T. J. Bronzert, and H. B. Brewer, Jr. 1972. Advances in the gas chromatographic analysis of amino acid phenyl- and methylthiohydantoins. Anal. Biochem. 45:43.
15 Ribadeau-Dumas, B., G. Brignon, F. Grosclaude, and J.-C. Mercier. 1972. Structure primairie de la casein β-bovine-Sequence complète. Eur. J. Biochem. 25:505.
16 Smithies, O., D. Gibson, E. M. Fanning, R. M. Goodfliesh, J. G. Gilman, and D. L. Ballantyne. 1971. Quantitative procedures for use with the Edman-Begg sequenator. Partial sequences of two unusual immunoglobulin light chains, Rzf and Sac. Biochemistry 10:4912.
17 Strid, L. 1961. Phosphopeptides from a tryptic hydrolysate of human casein. Acta Chem. Scand. 15:1423.
18 Thompson, M. P., and L. Pepper. 1964. Genetic polymorphism in caseins of cow's milk. IV. Isolation and properties of β-caseins A, B, and C. J.